

**WHAT IS CLAIMED IS:**

1. A method for characterizing a SAGE tag fragment comprising:
  - a) obtaining a RNA sample from the same tissue type as used in generating said SAGE tag;
  - b) generating cDNA fragments that correspond to the SAGE tag from said RNA sample by performing a DNA amplification reaction wherein primers used comprise:
    - (i) a SAGE tag sequence as a sense primer; and
    - (ii) at least one single-base anchored oligo-dT primer as an antisense primer; and
  - c) analyzing said cDNA fragments.
2. The method of claim 1, wherein said RNA sample is the RNA sample used to perform SAGE.
3. The method of claim 1, wherein said DNA amplification comprises polymerase chain reaction.
4. The method of claim 3, wherein the DNA polymerase used for said polymerase chain reaction is *Pfu* DNA polymerase.
5. The method of claim 3, wherein the  $Mg^{2+}$  concentration is 4 mM.
6. The method of claim 1, wherein said cDNA fragments generated are about 50 to 600 base pairs in length.
7. The method of claim 1, wherein said single-base anchored oligo-dT primer comprises a single-base anchored to the 3' end of the oligo-dT primer said base excluding dT.

8. The method of claim 1, wherein said single-base anchored oligo-dT primer comprises from 10 to 25 poly-dT residues.

5 9. The method of claim 8, wherein said single-base anchored oligo-dT primer is preferably comprised of 11 poly-dT residues.

10. The method of claim 1, wherein said sense primer further comprises a *Bam*HI recognition sequence at the 5' end.

10

11. The method of claim 1, wherein said SAGE tag further comprises a *Nla*III recognition sequence at the 5' end.

12. The method of claim 1, wherein said analyzing comprises:

15

- i) cloning said cDNA fragments; and
- ii) sequencing said clones to identify said cDNA fragment sequence.

13. The method of claim 12, further comprising comparing the cDNA sequence to sequences in existing DNA databases.

20

14. The method of claim 1, wherein said analyzing comprises hybridizing the cDNA fragments with known sequences.

25

15. The method of claim 1, wherein said analyzing comprises cloning the full-length cDNA.

16. The method of claim 1, wherein said analyzing comprises performing a DNA amplification reaction using:

30

- i) a sense primer designed based on an existing exon sequence;

- ii) a single-base anchored oligo-dT primer as an antisense primer; and
- iii) cloning and sequencing the amplified DNA.

17. The method of claim 16, wherein the exon sequences are predicted by  
5 bioinformatics tools.

18. The method of claim 17, further comprising aligning the sequence of the  
amplified cDNA with genomic DNA sequences.

19. The method of claim 1, wherein the tissue type is selected from the group  
consisting of colon, thymus, small intestine, heart, placenta, skeletal muscle, testes, bone  
marrow, trachea, spinal cord, liver, spleen, brain, lung, ovary, prostate, skin, cornea,  
retina, and breast.

20. The method of claim 15, wherein the full length cDNA is cloned into an  
expression vector.

21. A method for identifying a gene comprising:

- a) obtaining an isolated protein;
- b) digesting said protein to obtain at least a first protein fragment;
- c) obtaining at least a first amino acid sequence from said first protein  
fragment;
- d) generating a first DNA fragment that encodes said first protein fragment;
- e) performing a DNA amplification reaction with cDNA obtained from the  
same tissue sample as the isolated protein wherein primers used  
comprise:
  - (i) a sense primer comprising said first DNA; and
  - (iii) at least one single-base anchored oligo-dT primer as an antisense  
primer; and
- f) analyzing said cDNA fragments.

22. The method of claim 21, wherein the steps c) through f) are repeated with a second protein fragment.

23. The method of claim 21, wherein the steps c) through f) are repeated with a third protein fragment.

24. The method of claim 21, wherein the steps c) through f) are repeated with a fourth protein fragment.

25. The method of claim 21, wherein the steps c) through f) are repeated with a fifth protein fragment.

26. The method of claim 21, wherein said digesting protein is followed by separation to obtain digested protein fragments.

27. The method of claim 26, wherein said separation is based on the size of the protein fragments.

28. The method of claim 26, wherein said separation is by HPLC.

29. The method of claim 26, wherein said separation is by FPLC.

30. The method of claim 26, wherein said separation is by gel electrophoresis.

31. The method of claim 26, wherein said separation is by molecular sieve chromatography.

32. A method for characterizing a SAGE tag fragment comprising:

a) obtaining a RNA sample;

b) generating cDNA fragments using a 3' anchored oligo dT primer for first strand synthesis;

- c) digesting the cDNA generated in step b) with an enzyme;
- d) isolating 3' cDNA fragments of the digested cDNA;
- e) amplifying the 3' cDNA fragments of step d) by:
  - (i) ligating a SAGE linker to the 3' cDNA; and
  - (ii) mixing said 3' cDNA with a sense primer comprising the sequence of the SAGE linker, an antisense primer comprising the sequence of the primer used in step b) or a fragment thereof, and a polymerase enzyme, under conditions suitable for amplification;
- f) purifying the amplified 3' cDNA fragments obtained in step e);
- g) performing a second amplification comprising generation of longer cDNA fragments from SAGE tags in a multi-well format by mixing said cDNA fragments with a sense primer comprising a SAGE tag sequence and a restriction enzyme sequence, an antisense primer comprising the sequence of the primer used in step b) or a fragment thereof, and a polymerase enzyme, under conditions suitable for amplification;
- h) cloning and sequencing the products generated in step g).

33. The method of claim 32, wherein the 3' anchored oligo dT primer for first strand synthesis is further attached to an affinity label.

34. The method of claim 33, wherein the affinity label is biotin.

35. The method of claim 32, wherein the enzyme in step c) is a restriction enzyme.

36. The method of claim 35, wherein the enzyme is *Nla*III.

37. The method of claim 32, wherein the isolating comprises affinity-based isolating.

38. The method of claim 37, wherein the isolating utilizes streptavidin.

39. The method of claim 32, wherein the polymerase enzyme used in steps e) and g) is PLATINUM Taq.

40. The method of claim 32, wherein the cloning and sequencing comprises:

- a) precipitating and purifying the amplified products of step g) in the multi-well format;
- b) cloning the purified products into a vector;
- c) transforming competent bacteria with cloned products;
- d) screening for transformants; and
- e) sequencing DNA from transformants to identify the gene encoded by the SAGE tag.

41. The method of claim 32, wherein more than one SAGE tags are simultaneously identified.